



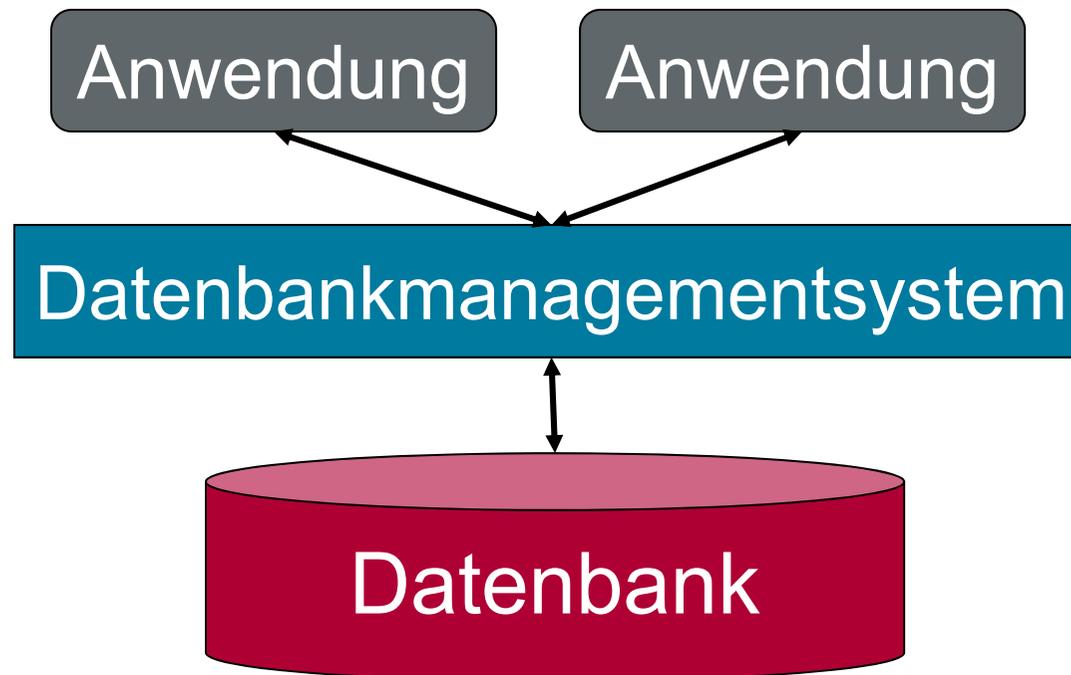
**Hasso
Plattner
Institut**

IT Systems Engineering | Universität Potsdam

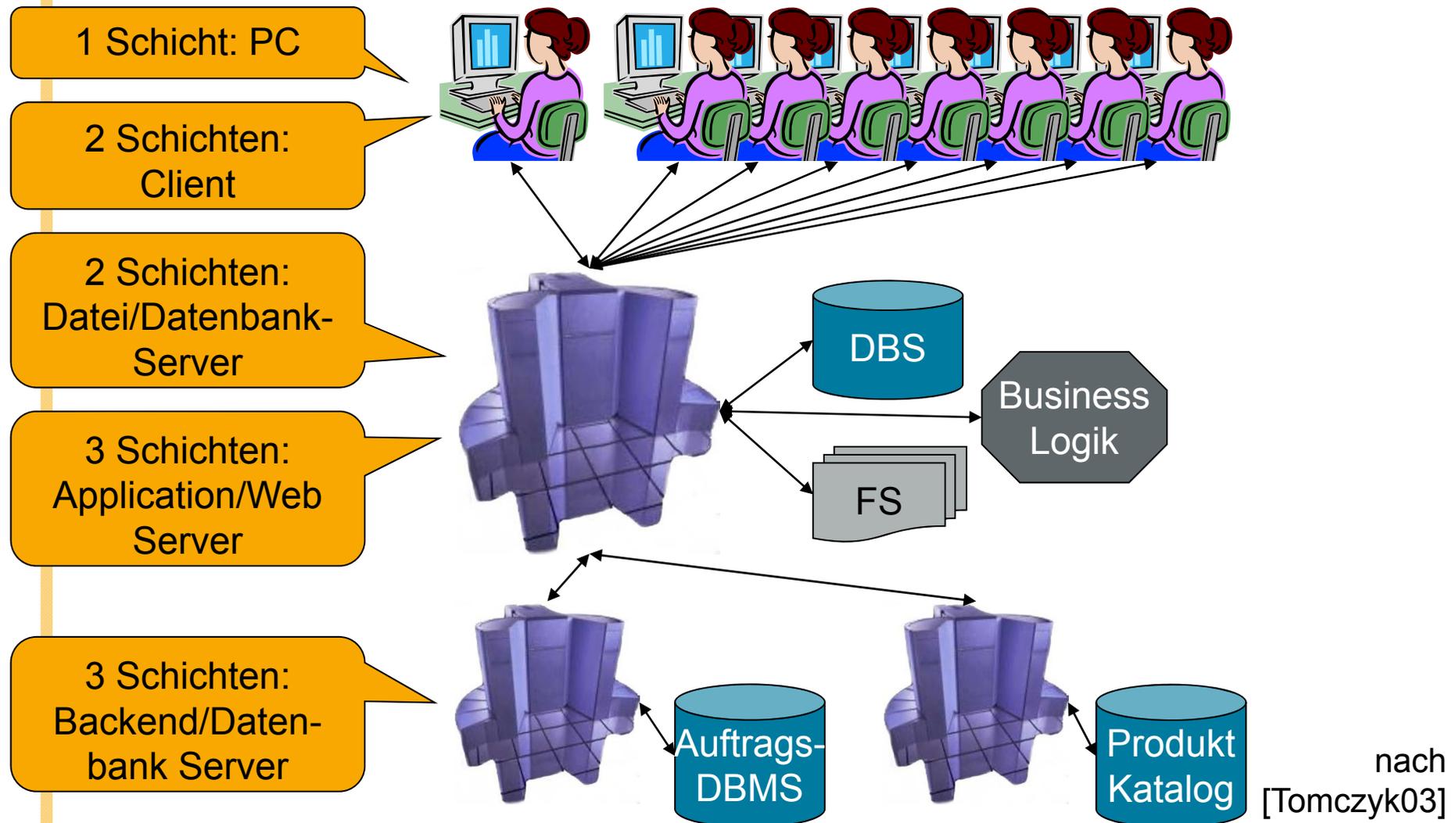
Datenbanksysteme 1
Organisatorisches und Einführung

19.4.2010

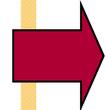
Felix Naumann



Mehr-Schichtenarchitekturen



4



- Vorstellung der Arbeitsgruppe
- Organisatorisches
- Datenbanken und Informationssysteme
- Das Semester an einem Beispiel
- Ausblick auf das Semester



Information Systems Team



Other courses in this semester

6

Lectures

- DBS I
- Search engines

Seminars

- Bachelor: Beauty is our Business
- Bachelor: No SQL
- Master: Collaborative Filtering
- Masterproject: Duplikaterkennung auf GPUs

Bachelorprojects

- LongCat: Data Profiling (IBM)
- Cathbad: Faceted Search (Excentos)



Extending the Database Relational Model to Capture More Meaning

E. F. CODD
IBM Research Laboratory

During the last three or four years several investigators have been exploring "semantic models" for formatted databases. The intent is to capture (in a more or less formal way) more of the meaning of the data so that database design can become more systematic and the database system itself can behave more intelligently. Two major thrusts are clear:

- (1) the search for meaningful units that are as small as possible—atomic semantics;
- (2) the search for meaningful units that are larger than the usual n -ary relation—molecular semantics.

In this paper we propose extensions to the relational model to support certain atomic and molecular semantics. These extensions represent a synthesis of many ideas from the published work in semantic modeling plus the introduction of new rules for insertion, update, and deletion, as well as new algebraic operators.

Key Words and Phrase: relation, relational database, relational model, relational schema, database, data model, database schema, data semantics, semantic model, knowledge representation, knowledge base, conceptual model, conceptual schema, entity model

CR Categories: 3.70, 3.73, 4.22, 4.25, 4.33, 4.34, 4.39

1. INTRODUCTION

The relational model for formatted databases [5] was conceived ten years ago, primarily as a tool to free users from the frustrations of having to deal with the clutter of storage representation details. This implementation independence coupled with the power of the algebraic operators on n -ary relations and the open questions concerning dependencies (functional, multivalued, and join) within and between relations have stimulated research in database management (see [30]). The relational model has also provided an architectural focus for the design of databases and some general-purpose database management systems such as MACAIMS [13], PRTV [38], RDMS(GM) [41], MAGNUM [19], INGRES [37], QBE [46], and System R [2].

During the last few years numerous investigations have been aimed at capturing

Permission to copy without fee all or part of this material is granted provided that the copies are not made or distributed for direct commercial advantage, the ACM copyright notice and the title of the publication and its date appear, and notice is given that copying is by permission of the Association for Computing Machinery. To copy otherwise, or to republish, requires a fee and/or specific permission.

A version of this work was presented at the 1979 International Conference on Management of Data (SIGMOD), Boston, Mass., May 30–June 1, 1979.
Author's address: IBM Research Laboratory K01/282, 5600 Cottle Road, San Jose, CA 95193.
© 1979 ACM 0362-5915/79/1200-0387 \$00.75

ACM Transactions on Database Systems, Vol. 4, No. 4, December 1979, Pages 387–434.

HOW TO WRITE A CV



Leverage the NoSQL boom

Bachelorseminar NoSQL

8

- Get to know a specific NoSQL system in theory and practice
- learn about and perform database benchmarking
 - [Yahoo! Cloud Serving Benchmark](#)
- Theoretical part
 - choose a specific NoSQL system and gather information about it
 - describe the architecture and implementation of the system
 - discuss differences between the system and traditional relational databases
 - discuss differences between the system and other NoSQL databases
- Hands-on part
 - Install the NoSQL database
 - Install [Yahoo! Cloud Serving Benchmark](#)
 - Add connector to the YCSB
 - Benchmark the NoSQL system
- NoSQL systems
- Wide Column/Key-Value Stores
 - [Apache Cassandra](#)
 - [Apache HBase](#)
 - [Tokyo Cabinet](#)
 - [Redis](#)
 - [Riak](#)
- Document Stores
 - [CouchDB](#)
 - [MongoDB](#)
- Graph Databases
 - [neo4j](#)
 - [OrientDB](#)
 - [InfoGrid](#)

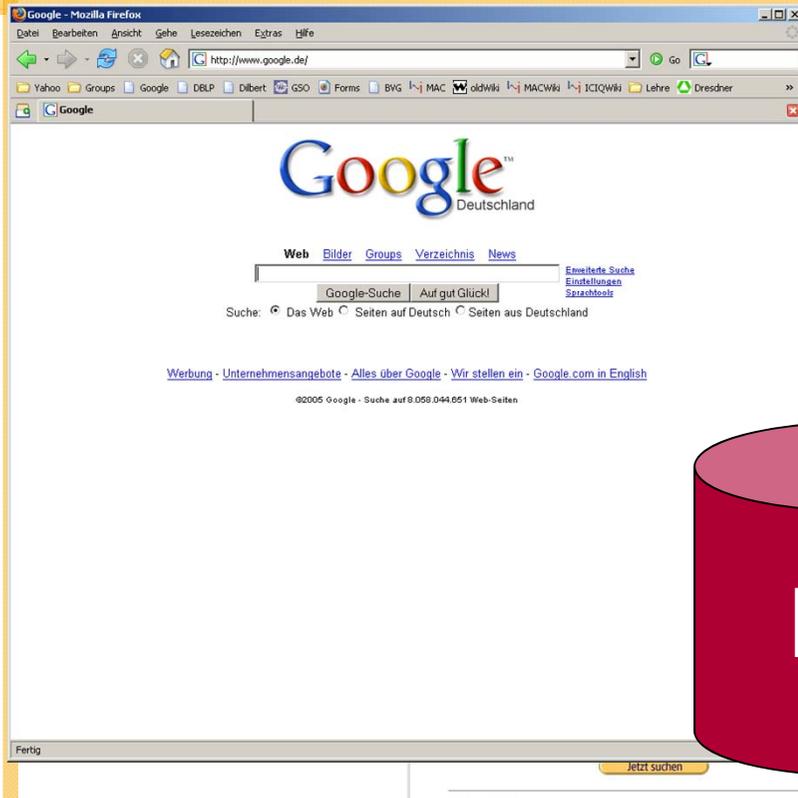
Beauty is our Business

„Wenn wir uns klarmachen, daß der Kampf gegen Chaos, Durcheinander und unbeherrschte Kompliziertheit eine der größten Herausforderungen der Informatik ist, müssen wir zugestehen: Beauty is our Business.“ Edsger W. Dijkstra, 1978

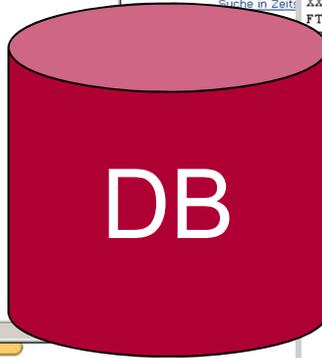
Termin	Thema	Vortrag
11.4.2011	Einführung	Felix Naumann
18.4.2011	Wissenschaftliche Texte Lesen	Felix Naumann
9.5.2011	Elevatorpitches / Diskussion	Alle
16.5.2011	Vortragstechniken	Felix Naumann
13.6.2011	Vortrag 1: Mariposa (pdf) Vortrag 2: Trio (pdf)	
20.6.2011	Vortrag 3: Fagins Algorithmus (pdf) Vortrag 4: Enough Already in SQL (pdf)	
27.6.2011	Vortrag 5: Sorted Neighborhood (pdf) Vortrag 6: Data Mining (pdf)	
4.7.2011	Einführung in LaTeX	Felix Naumann
18.7.2011	Vorstellung der Gliederungen & Tipps zur Ausarbeitung	Alle
tbd	Abgabe der Ausarbeitungen	Alle

Was sind Informationssysteme?

10

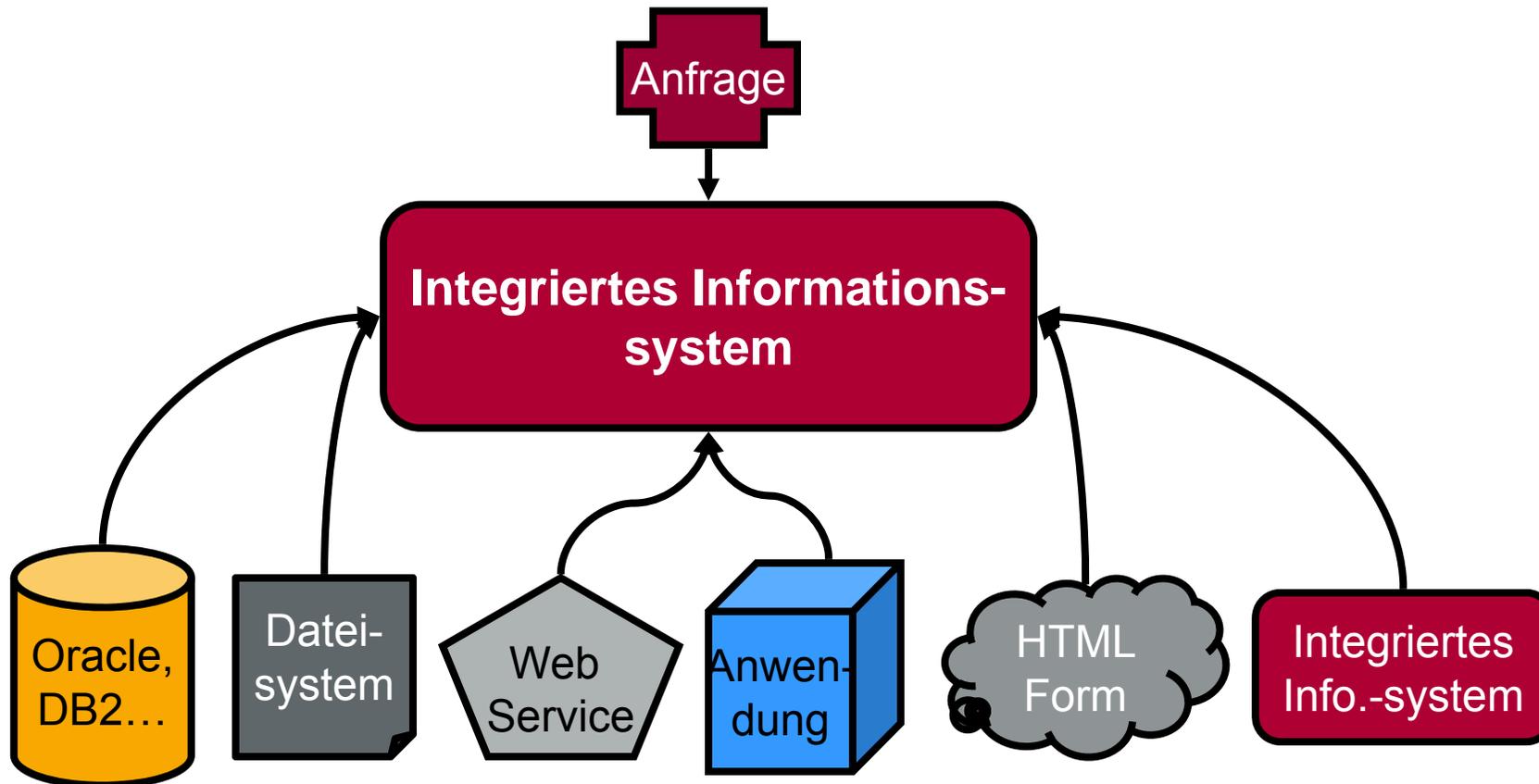


ID	RNGTPCHI	standard; RNA; ROD: 1016 BP.	Molecule type
XX			Name
DT	01-AUG-1991 (Rel. 28, Created)		Date of creation and last update
DT	04-MAR-2000 (Rel. 63, Last updated, Version 2)		
XX			
DE	Rat GTP cyclohydrolase I mRNA, complete cds.		Free text description
XX			
KW	GTP cyclohydrolase I.		Keywords describing the molecule
XX			
OS	Rattus norvegicus (Norway rat)		Organism
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;		
OC	Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.		
XX			
RN	[1]		Article the sequence was published in
RP	1-1016		
RX	MEDLINE; 91093270.		
RX	PUBMED; 1985963.		
RA	Hatakeyama K., Inoue Y., Harada T., Kagamiyama H.;		
RT	"Cloning and sequencing of cDNA encoding rat GTP cyclohydrolase I: The first enzyme of the tetrahydrobiopterin biosynthetic pathway";		
RL	J. Biol. Chem. 266(2):765-769(1991).		
XX			
FT	CDS	128..853	Structural annotation (coding sequence)
		/codon_start=1	
		/db_xref="GOA:P22288"	Link to functional annotation of resulting protein
		/db_xref="SWISS-PROT:P22288"	
		/EC_number="3.5.4.16"	
		/gene="GTP cyclohydrolase I"	
		/product="GTP cyclohydrolase I"	
		/protein_id="AA41299.1"	
		/translation="MEKPRGVRCTNGFPERELPRFGASRPAEKSRPPEAKGQPADAWK	Translated protein sequence
		AGRPRSEEDNELNLPNLAAAYSSLRLSGEDPQRQGLLKTTPURAAATAMQFFTKGYQETI	
		SDVLND&IFDEHDHDMVIVKIDIMFSMCEHHLVFPVGRVHIGVLPNKQVLGSLKLARIV	
		EIYSRRLQVQERLTKQIAVAITEALQPAAGVGVVIEATHMCMVNRGVQKMNSTVTSITML	
		GVFREDPKTREFLTLIRSI"	
		Sequence 1016 BP; 236 A; 279 C; 291 G; 210 T; 0 other:	
		gacttcgaac ctcattcgtt gcagaactcc tgtcccgggt acagccacag gtcacggcgc	60
		ccggctaagc cgagccgcag cgtttgtag caacttaggt tgtctcggga gcaatcgcgc	120
		cggtccatg gagaagccgc ggggtgtaag gtgcaccaat ggggtccccg agcgggagct	180
		...	
		catcaggagc tgaactccg tgtgcgagcc ccggtttgca gacccccgt gaggccagcg	900
		ttatctgtct cgattgtaca ttccagttcc agttggata ctgtgcaact ttattctca	960
		ccatgaattg tattaataa ttatttatag agatgtcaaa taaaggtgat caactt	1016



```
<buch>
  <isbn>0-201-318051</isbn>
  <titel>XML und Datenbanken</titel>
  <autor>Klettke/Meyer</autor>
</buch>
```





Schematische und Daten-Heterogenität

12

Variante 1

Männer	
Vorname	Nachname
Felix	Naumann
Jens	Bleiholder

Frauen	
Vorname	Nachname
Melanie	Weis
Jana	Bauckmann

Variante 2

Personen			
Vorname	Nachname	Männl.	Weibl.
Felix	Naumann	Ja	Nein
Jens	Bleiholder	Ja	Nein
Melanie	Weis	Nein	Ja
Jana	Bauckmann	Nein	Ja

Variante 3

Personen		
Vorname	Nachname	Geschlecht
Felix	Naumann	Männlich
Jens	Bleiholder	Männlich
Melanie	Weis	Weiblich
Jana	Bauckmann	Weiblich

Schematische und Daten-Heterogenität

13

Variante 1

Männer	
Vorname	Nachname
Felix	Naumann
Jens	Bleiholder

Frauen

Vorname	Nachname
Melanie	Weis
Jana	Bauckmann

Variante 2

Personen			
FirstNa	Name	male	femal
Felix	Naumann	Ja	Nein
Jnes	Bleiho.	Ja	Nein
Melanie	Weiß	Nein	Ja
Jana	baukman	Nein	Ja

Variante 3

Personen		
VN	NN	SEX
F.	Naumann	Männlich
J.	Bleiholder	Männlich
M.	Weis	Weiblich
J.	Bauckmann	Weiblich

Schematische und Daten-Heterogenität

14

Variante 1

*■□■□■	
†□○■	♣■□○■
♣■●×	♣○◆○■
☺■●	×■□●○■

♣○◆■	
†□○■	♣■□○■
♣■●○	♣■×
☺■	♣&○■

Variante 2

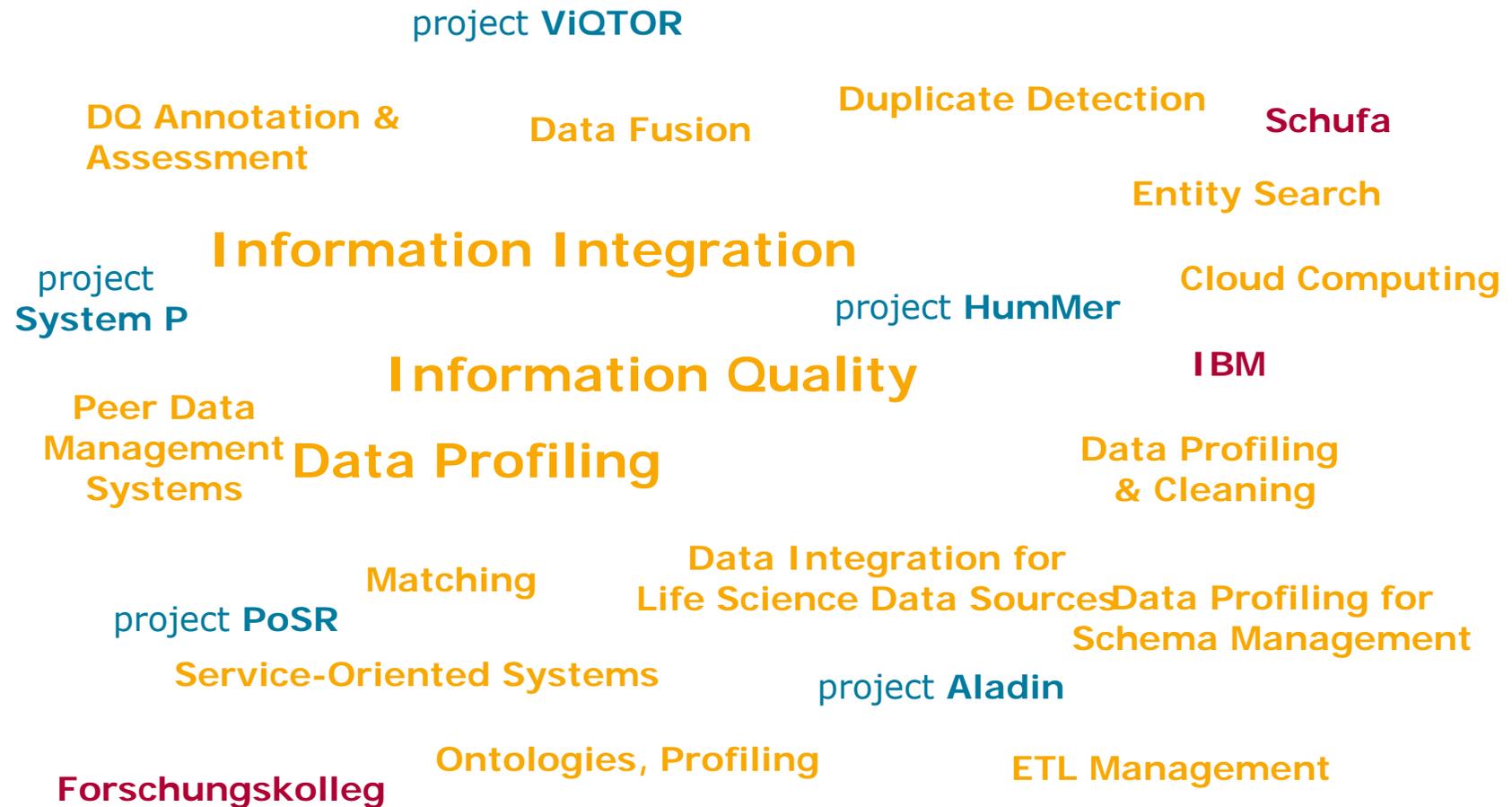
♣■□○■			
†□□■	♣■○■	■●	●
♣■●×	♣○◆○■	📁	📁
☺■●	♣●○■	📁	📁
♣■●○	♣■×	📁	📁
☺■	&○■	📁	📁

Variante 3

♣■□○■		
■○■	♣■○■	♣●■×◆
♣■●×	♣○◆○	♣■●×♣
☺■●	♣●○■	♣■●×♣
♣■●○	♣■×	♣○●×♣
☺■	♣○◆×	♣○●×♣

Weitere Forschungsthemen

15



Vorstellung – Hörer

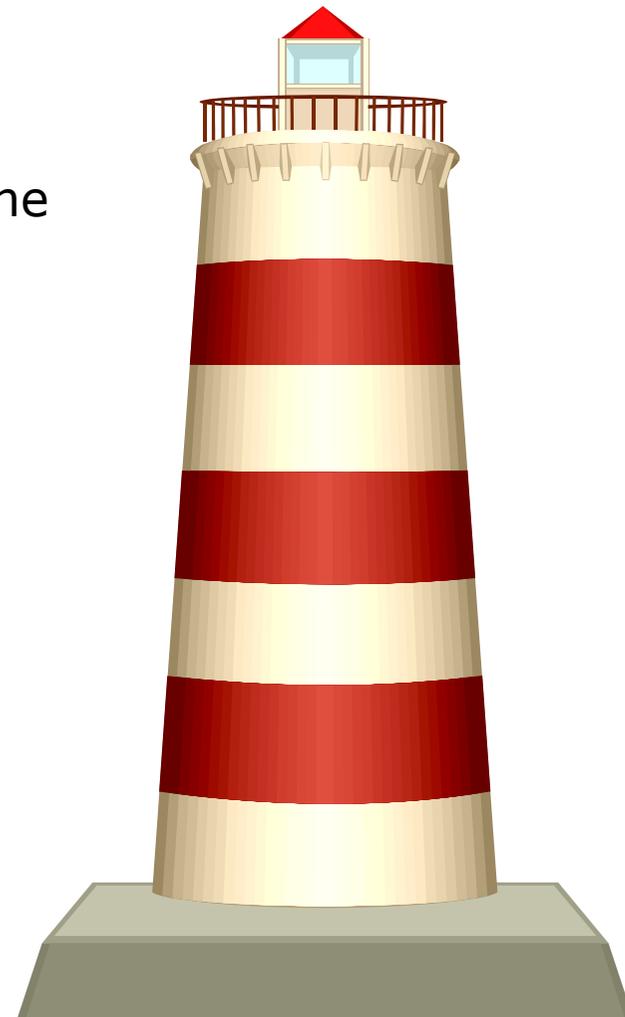
16

- Welches Semester?
 - 2. bzw 4. Semester
- Erasmus o.ä.?
- Berlin/Brandenburg oder anderswo?

- Berührung mit Datenbanken?
 - Wie?

17

- Vorstellung der Arbeitsgruppe
- Organisatorisches
- Datenbanken und Informationssysteme
- Das Semester an einem Beispiel
- Ausblick auf das Semester



Termine

18

- Montags 9:15 - 10:45
 - Hörsaal 1
- Mittwochs 13:30 - 15:00 (hungrig)
 - Hörsaal 2
- Erste Vorlesung
 - 11.4.2011
- Letzte Vorlesung
 - 20.7.2011
- Relevante Feiertag(e)
 - Ostermontag am 25. April
 - Pfingstmontag am 13. Juni
- Klausur
 - Vielleicht letzte Semesterwoche
 - 180(!) Minuten
- Übungen
 - Ungefähr 1 SWS
 - Verteilt auf Doppelstunden im Semester
 - Drei parallele Übungsgruppen
- Übungsleitung: Jana Bauckmann
- Tutoren-Team
 - Fabian Tschirschnitz
 - Claudia Exeler
 - Jan-Peer Rudolph

Übung: Termine

- 3 identische Übungstermine pro Woche
 - Mittwoch, 13:30-15:00 Uhr in HS2 (anstelle der Vorlesung)
 - Freitag, 9:15-10:45 Uhr in A-1.2
 - Freitag, 11:00-12:30 Uhr in A-1.2

- Anmeldung
 - Email – bis 17.4. – an dbs1-2011@hpi.uni-potsdam.de mit Wunschliste der drei Termine
 - Wir ordnen die Übungstermine zu und veröffentlichen das Ergebnis am 19.4. unter
\\fs3\lehrveranstaltungen\FG_Informationssysteme\DBS I\Uebung2011

Übung: Aufgabenblätter

- insgesamt 7 Aufgabenblätter (aber nur 6 Übungstermine)
 - „Theoretische“ Übungen auf Papier
 - Praktische Übungen
 - ◇ VMware-Image mit DB2, VMware Player in den Poolräumen (oder auf eigenem Rechner)
 - ◇ Ergebnisse auf Papier
 - Erstes Aufgabenblatt ohne Übungstermin!
 - ◇ DBMS-Historie
 - ◇ Aufgabenblatt ab 18.4. auf Webseite, Abgabe 26.4.
- Bearbeitung der Aufgabenblätter in **2er Gruppen**

Übung: Aufgabenblätter (cont.)

- Prüfungszulassung
 - Maximal ein Aufgabenblatt mit mind. 25% der Punkte bewertet
 - Alle weiteren Aufgabenblätter mit mind. 50% der Punkte bewertet

- Abgabe
 - Abgabetermin: Dienstag nach der Übungswoche
 - per Email an dbs1-2009@hpi.uni-potsdam.de
 - pro Aufgabe eine Datei mit Dateiname:
blatt<aufgabenblattNr>aufgabe<aufgabenNr> <Namen>.pdf

Übung: Hilfe bei Problemen

- Mailingliste für Fragen zur Übung: fragen-dbs1-2011@hpi...
 - Hilfe zur Selbsthilfe: Fragen gegenseitig beantworten
 - Mitarbeiter lesen (und antworten) auch mit
 - wir tragen alle ein, die ihre Übungsterminwünsche an dbs1-2011@hpi... senden
oder anmelden unter <http://lists.hpi.uni-potsdam.de/listinfo/fragen-dbs1-2011>

- Sprechstunde Jana (in den Übungswochen)
 - Freitag 12:30 Uhr in A-1.2 (nach der Übung)

- Folien
 - Vor der VL im WWW
 - ASAP
- Anregungen zur Verbesserung:
 - Gebrauch der Folien
 - Infos im WWW
 - Übungsbetrieb
 - ...
- Fragen bitte jederzeit!
 - In der Vorlesung
 - Sprechstunde
 - ◇ Dienstags 14:00 – 15:00
 - ◇ Raum A-1.13
 - ◇ Am liebsten mit Anmeldung
 - Email: naumann@hpi.uni-potsdam.de

Feedback

24

- gut gelaunt und engagiert
- vermittelt positive Stimmung
- aber redet ein wenig zu schnell und zu viel

Die Bewertung bzw. Korrigierung der Übungsaufgaben ist sehr schön, durch das genaue Korrigieren erhält man einen guten Feedback.

Es ist beruhigend, wenn man merkt, wenn der Dozent bei einem bestimmten Problem auch nicht 100% sicher ist.

Vote
Naumann for
Best-Prof-
WS-06/07

Die Vorlesung war sehr angenehm, auch theoretische Teile wurden so dargestellt, dass man nicht gleich eingepennt ist.

Bitte kein DB2
mehr! Bitte bitte!

Dozent könnte
sein Sprechtempo
ab und zu etwas
herabsetzen

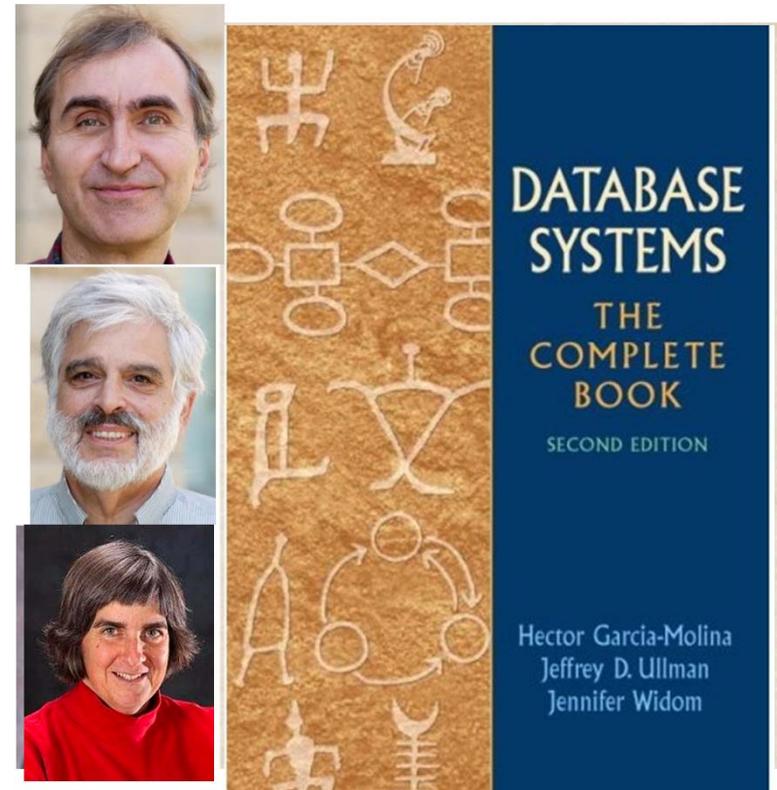
Es wäre schön wenn Sie in
Zukunft langsamer sprechen
könnten bzw. manchmal eine
Kunstpause einschieben könnten

Database Systems - The Complete Book

- Hector Garcia-Molina, Jeffrey D. Ullman, Jennifer Widom: Pearson Education International, 2002.

Ebenfalls empfehlenswert

- Datenbanksysteme
Alfons Kemper, André Eickler
ISBN: 3486273922
- Datenbanken:
Implementierungstechniken
Gunter Saake, Andreas Heuer, Kai-
Uwe Sattler



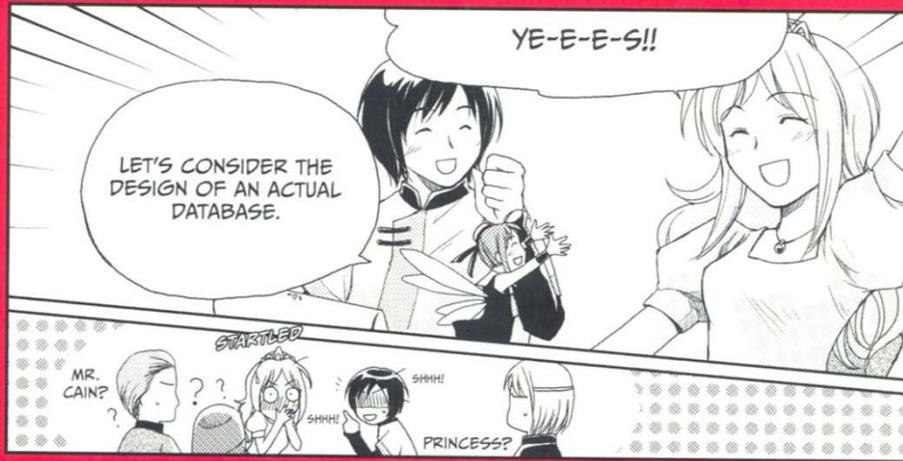
Und viele andere mehr...

Weitere Literatur

UNLEASH THE POWER OF THE DATABASE!

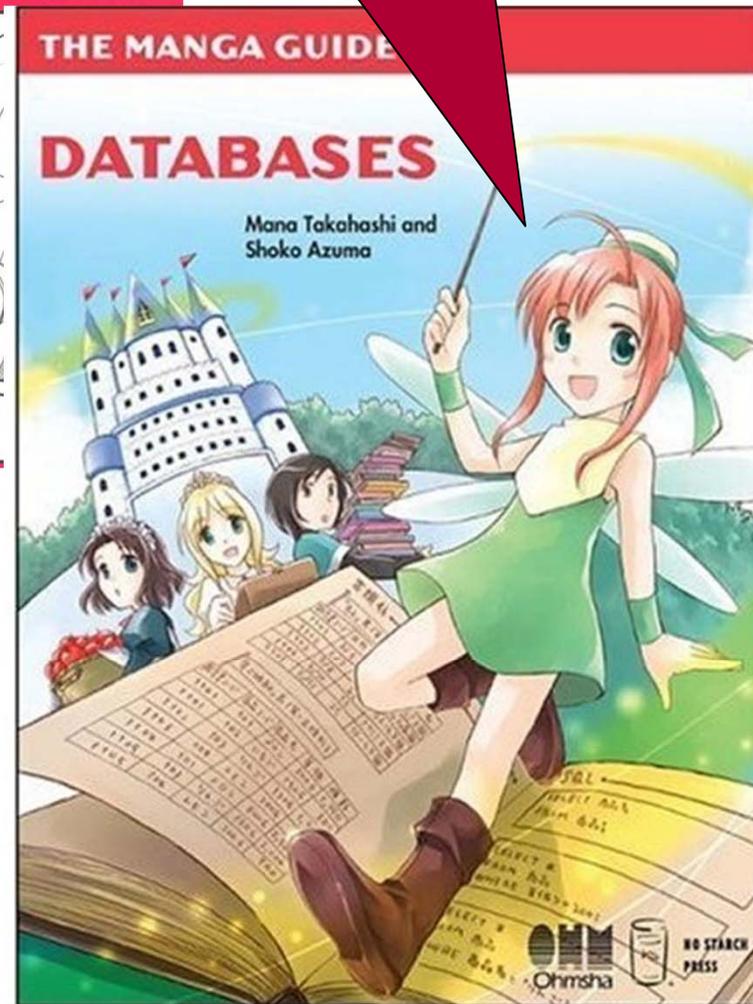
Magical database fairy

26



PRINCESS RURUNA AND CAIN HAVE A PROBLEM: THEIR FRUIT-SELLING EMPIRE IS A TANGLE OF CONFLICTING AND DUPLICATED DATA, AND SORTING THE MELONS FROM THE APPLES AND STRAWBERRIES IS CAUSING REAL DIFFICULTIES. BUT WHAT CAN THEY DO?

WHY, BUILD A RELATIONAL DATABASE OF COURSE, WITH THE HELP OF TICO, THE MAGICAL DATABASE FAIRY. FOLLOW ALONG IN *THE MANGA GUIDE TO DATABASES* AS TICO TEACHES RURUNA AND CAIN HOW TO BUILD A DATABASE TO MANAGE THEIR KINGDOM'S SALES, MERCHANDISE, AND EXPORTS. YOU'LL LEARN HOW DATABASES WORK AND THE MEANING OF TERMS LIKE SCHEMAS, KEYS, NORMALIZATION, AND TRANSACTIONS.



- Vorstellung der Arbeitsgruppe
- Organisatorisches
- Datenbanken und Informationssysteme
- Das Semester an einem Beispiel
- Ausblick auf das Semester



Was sind Daten?

28

Digitale Repräsentation von

- Dingen
- Entitäten
- Wissen
- Information

in/aus der wirklichen Welt.

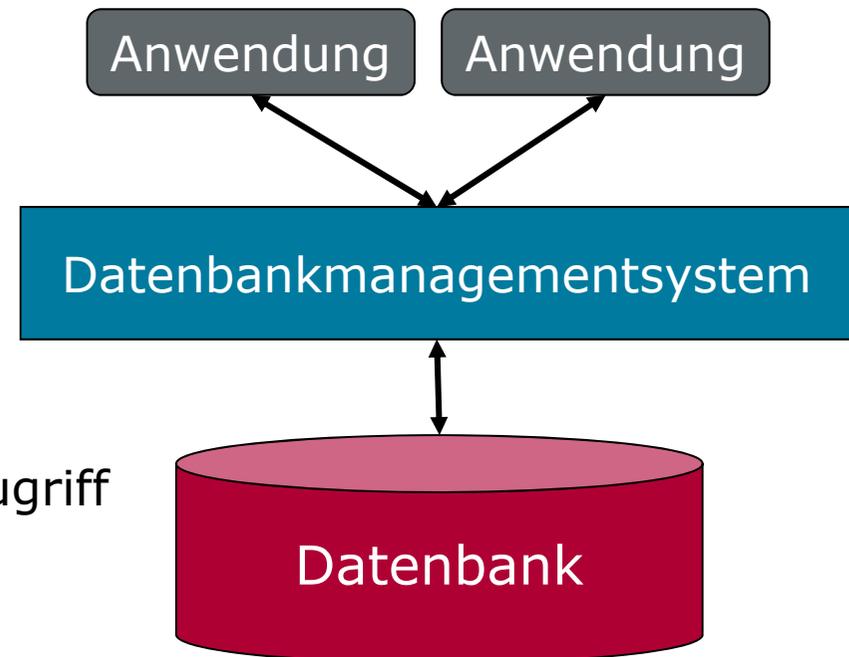
Kernfragen:

- Welche Daten speichere ich?
- Wie speichere ich die Daten?
- Wie frage ich Daten ab?
- Wie erledige ich all dies effizient und sicher?

} Unterstützung durch
Datenbanksystem

Bestandteile

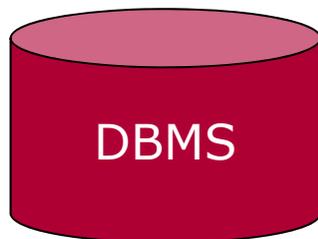
- Datenbank
 - Die Daten selbst
 - + Metadaten (Beschreibung der Daten)
- Datenbankmanagementsystem (DBMS)
 - Softwarekomponente zum Zugriff auf eine oder mehrere Datenbanken.
 - Server-basiert
- (Anwendungen)



DBMS Beispiele

30

- DBMS
 - OLTP (Online Transaction Processing)
 - Oracle, DB2, SQL-Server
 - Informix, Sybase
 - Teradata
 - PostgreSQL, Interbase
 - MySQL, ...
- Data Warehouses
 - OLAP (Online Analytical Processing)
- Alle großen Softwaresysteme nutzen ein, mehrere oder sehr viele DBMS.
 - SAP
 - Siebel
 - SABRE
 - Banken, Versicherungen, Telekoms, usw.



ORACLE

tamino

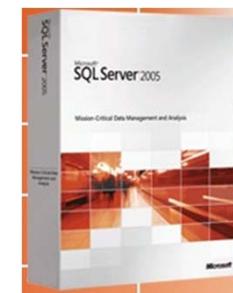
SYBASE



DB2 9 for Linux UNIX and Windows

pureXML™ and storage compression

Information Management software



DBMS Aufgaben

31

- Unterstützung des Datenmodells
- Bereitstellung einer Anfragesprache (DDL & DML)
- Effiziente Anfragebearbeitung
- Robustheit
 - Wahrung der Datenintegrität (Konsistenz etc.)
 - Abfangen von Systemfehlern
- Speicherverwaltung (RAM & Disk)
- Transaktionsmanagement
 - Auch im Mehr-Benutzer-Betrieb
- Nutzerverwaltung & Zugangskontrolle

Beispiele für Informationssysteme

33

Datei

- Informationseinheit: Zeile /Token

- Anfrage: Parser

- Struktur: Flach

Beispiele

- Komma-delimited files
- Annotated files

Einsatzgebiete

- SwissProt

```

ID  RNGTPCHI  standard; RNA; ROD; 1016 BP.
XX
DT  01-AUG-1991 (Rel. 28, Created)
DT  04-MAR-2000 (Rel. 63, Last updated, Version 2)
XX
DE  Rat GTP cyclohydrolase I mRNA, complete cds.
XX
KW  GTP cyclohydrolase I.
XX
OS  Rattus norvegicus (Norway rat)
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;
OC  Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
XX
RN  [1]
RP  1-1016
RX  MEDLINE; 91093270.
RX  PUBMED; 1985963.
RA  Hatakeyama K., Inoue Y., Harada T., Kagamiyama H.;
RT  "Cloning and sequencing of cDNA encoding rat GTP cyclohydrolase I: The
RT  first enzyme of the tetrahydrobiopterin biosynthetic pathway";
RL  J. Biol. Chem. 266(2):765-769(1991).
XX
FT  CDS 128..853
FT  /codon_start=1
FT  /db_xref="GOA:P22288"
FT  /db_xref="SWISS-PROT:P22288"
FT  /EC_number="3.5.4.16"
FT  /gene="GTP cyclohydrolase I"
FT  /product="GTP cyclohydrolase I"
FT  /protein_id="AAA41299.1"
FT  /translation="MEKPRGVRCNTNGFPERELPRPGASRPAEKSRPEAKGAQPADAWK
FT  AGRPRSEEDNELNLPNLAAAYSSILRSLGEDPQRQLLKTWPRAATAMQFFTRGYQETI
FT  SDVLNDAIFDEHDENVIVKIDIDMFSCHEHLLVPFVGRVHIGYLPNKQVLGSLKLARIV
FT  EIYSRRLQVQERLTKQIAVAITEALQFAGVGVVIEATHMCMVHRGVQKMNKTVTSTML
FT  GVFREDPKTREEFLTLIRS"
SQ  Sequence 1016 BP; 236 A; 279 C; 291 G; 210 T; 0 other;
gacttcgaac ctcattcggg gcagaactcc tgtcccgggtg acagccacag gtcacggcgg 60
cgggctaagc cgagccgcag cgcttggtag caccttaggg tgtctcgga gcaatcgcgc 120
cgggtccatg gagaagccgc ggggtgtaag gtcaccaat gggttccccg agcgggagct 180
...
catcaggagc tgaacttcgg tgtgcgagcc cgggtttgca gacccccgct gaggccagcg 900
ttatctgtct cgattgtaca ttccagttcc agttgggata cttgtcaact ttatttctca 960
ccatgaattg tattaataa ttatttatag agatgtcaaa taaaggtgat caactt 1016
//

```

Molecule type
Name
Date of creation and last update
Free text description
Keywords describing the molecule
Organism
Article the sequence was published in
Structural annotation (coding sequence)
Link to functional annotation of resulting protein
Translated protein sequence
Sequence of bases

Beispiel: XML-Dokument

```
35 <?xml version="1.0" encoding="UTF-8"?>
<rechnung kundennummer="k333063143">
  <monatspreis>0,00</monatspreis>
  <einzelverbindungsachweis>
    <verbindung>
      <datum>26.2.</datum>
      <zeit>19:47</zeit>
      <nummer>200xxxx</nummer>
      <einzelpreis waehrung="Euro">0,66</einzelpreis>
    </verbindung>
    <verbindung>
      <datum>27.2.</datum>
      <zeit>19:06</zeit>
      <nummer>200xxxx</nummer>
      <einzelpreis waehrung="Euro">0.46</einzelpreis>
    </verbindung>
    <verbindungskosten_gesamt waehrung="Euro">2.19</verbindungskosten_gesamt>
  </einzelverbindungsachweis>
</rechnung>
```

Beispiel: HTML Dokument

36

Google **squared** labs Unsave

Item Name	Image	Description	Authors	Page Count	Publisher	Product Form	Publicat
<input checked="" type="checkbox"/> Database systems: design, implementation, and management		DATABASE SYSTEMS: DESIGN, IMPLEMENTATION, AND MANAGEMENT, NINTH EDITION, a market-leader for database texts, gives readers a solid	Carlos Coronel, Steven Morris, Peter Rob	692	COURSE TECHNOLOGY	paperback	2009.11.2
<input checked="" type="checkbox"/> Database systems: a practical approach to design, implementation, and management		This best-selling text introduces the theory behind databases in a concise yet comprehensive manner, providing database design methodology that can be used by	Thomas M. Connolly, Carolyn E. Begg	1243	Addison-Wesley	paperback	2010
<input checked="" type="checkbox"/> Fundamentals of Database Systems		Clear explanations of theory and design, broad coverage of models and real systems, and an up-to-date introduction to modern database technologies result in a leading	Ramez Elmasri, Shamkant B. Navathe	1172	Addison-Wesley	hardback	2010.03.3
<input checked="" type="checkbox"/> An introduction to database systems		Continuing in the eighth edition, An Introduction to Database Systems provides a comprehensive introduction to the now very large field of database systems by providing	C. J. Date	1005	Pearson/Addison Wesley	paperback	2004
<input checked="" type="checkbox"/> Active rules in database systems		A timely survey of database methodology, this volume is organized by theme and is from the point of view of some of the subject's most active researchers. It covers	Norman W. Paton	439	Springer	hardback	1999
<input checked="" type="checkbox"/> A first course in database systems		Written by well-known computer scientists, this accessible and succinct introduction to database systems focuses on database design and use. It provides in-depth coverage	Jeffrey D. Ullman, Jennifer Widom	565	Pearson Prentice Hall	hardback	2008
<input checked="" type="checkbox"/> Database systems: the complete book		This introduction to database systems offers a comprehensive approach, focusing on database design, database use, and implementation of database applications and	Hector Garcia-Molina, Jeffrey D. Ullman, Jennifer Widom	1203	Pearson Prentice Hall	hardback	2009
<input checked="" type="checkbox"/> Spatial database systems: design, implementation and project management		This first of a kind book places spatial data within the broader domain of information technology (IT) while providing a comprehensive and coherent explanation of	Albert K. W. Yeung, G. Brent Hall	553	Springer	paperback	2007
<input checked="" type="checkbox"/> Encyclopedia of Database Systems		The multi-volume Encyclopedia of Database Systems provides easy access to relevant information on all aspects of very large databases, data management, and database	M. Tamer Özsu, Ling Liu	3749	Springer	hardback	2009.05.2
<input checked="" type="checkbox"/> Advances in Intelligent Information and Database Systems		Based on submissions to the 2nd Asian Conference on Intelligent Information and Database Systems, this proceedings covers the integration of artificial intelligence and	Ngoc Thanh Nguyen, Radosław Katarzyniak, Shyi-Ming Chen	370	Springer	hardback	2010.04.0
<input checked="" type="checkbox"/> Advances in Visual Information Management:		This state-of-the-art book explores new concepts, tools, and techniques for both	Hiroshi Arisawa, Tiziana Catarci	410	Kluwer Academic Publishers	hardback	2000.04.3

Beispiel: HTML Dokument

37



¹ Sieg nach Verlängerung

² Sieg im Elfmeterschießen

Achtelfinale [\[Bearbeiten\]](#)

→ *Hauptartikel: Fußball-Weltmeisterschaft 2010/Finalrunde#Achtelfinale*

Nr.	Tag und Uhrzeit	Spielort	Erste Mannschaft	Zweite Mannschaft	Ergebnis
1	Sa., 26. Juni 2010, 16:00 Uhr	Port Elizabeth	Uruguay	- Südkorea	2:1 (1:0)
2	Sa., 26. Juni 2010, 20:30 Uhr	Rustenburg	Vereinigte Staaten	- Ghana	1:2 n. V. (1:1, 0:1)
3	So., 27. Juni 2010, 16:00 Uhr	Bloemfontein	Deutschland	- England	4:1 (2:1)
4	So., 27. Juni 2010, 20:30 Uhr	Johannesburg (Soccer City)	Argentinien	- Mexiko	3:1 (2:0)
5	Mo., 28. Juni 2010, 16:00 Uhr	Durban	Niederlande	- Slowakei	2:1 (1:0)
6	Mo., 28. Juni 2010, 20:30 Uhr	Johannesburg (Ellis Park)	Brasilien	- Chile	3:0 (2:0)
7	Di., 29. Juni 2010, 16:00 Uhr	Pretoria	Paraguay	- Japan	0:0 n. V., 5:3 i. E.
8	Di., 29. Juni 2010, 20:30 Uhr	Kapstadt	Spanien	- Portugal	1:0 (0:0)

Viertelfinale [\[Bearbeiten\]](#)

→ *Hauptartikel: Fußball-Weltmeisterschaft 2010/Finalrunde#Viertelfinale*

Nr.	Tag und Uhrzeit	Spielort	Erste Mannschaft	Zweite Mannschaft	Ergebnis
1	Fr., 2. Juli 2010, 16:00 Uhr	Port Elizabeth	Niederlande	- Brasilien	2:1 (0:1)

Beispiel DBMS

39

Beispiel: HTML Formular

Suche Bücher

LOS

HARRY POTTER
und die Heiligtümer des Todes

Frühstück mit Harry [Jetzt portofrei vorbestellen](#)
am 27.10.2007 bis 10:30 Uhr geliefert

Erweiterte Suche Bücher

Je mehr Felder Sie ausfüllen, desto zielgerichteter können wir suchen. Es reicht jedoch aus, nur eines der Felder auszufüllen.

Autor/in:

Titel:

Schlagwörter:

ISBN: (10- oder 13-stellig, ohne Bindestriche)

Verlag:

Verfeinern Sie Ihre Suche, indem Sie nur nach bestimmten Buchformaten suchen lassen.

Nur gebraucht:

Format: Alle Formate

Ordnen nach: Topseller

Erscheinungsdatum: vor

Jeder Monat

2009

Suche in: deutsche Bücher englische Bücher

Jetzt suchen

Dahinter: Fast immer ein DBMS!

Beispiel: Web Services

43

X METHODS

[Home](#) · [Tools](#) · [Implementations](#) · [Manage](#) · [Register](#) · [Tutorials](#) · [About](#)



Full Service List Also accessible via XML Interfaces: [DISCO](#) [WS-Inspection](#) [RSS](#) See the [interfaces section](#) for more information.

List is ordered by submission time, with most recent services listed first.

Publisher	Style	Service Name	Description	Implementation
VOORSPRONG	DOC	Try It Worldcup 2010 Football Championships	Access information such as games, cities, results, players etc. about the 2010 tournament	Visual Dataflex
rpamplona	DOC	Try It ACHWorks-SOAP Ver 4(TT\$ - Rico Pamplona)	Web Services for ACH Processing and Payments	MS .NET
nsreevatsan	RPC	Try It Chennai Emergency	Emergency Service Provider	Apache SOAP
RestFulServices	DOC	Try It Weather Forecast Service	Provide weather forecast for a given place name	
RestFulServices	DOC	Try It HCPCS Service	The Healthcare Common Procedure Coding System (HCPCS) is a set of health care procedure codes based on the American Medical Association's Current Procedural Terminology (CPT). Commonly pronounced Hick-Picks.	
RestFulServices	DOC	Try It UNSPSC Service	UNSPSC is the acronym for the United Nations Standard Products and Services Code. UNSPSC is a coding system to classify both products and services for use throughout the global eCommerce marketplace.	
RestFulServices	DOC	Try It London Gold Fix Service	The Gold Fixing (also known as the London Gold Fixing or Gold Fix) is the procedure by which the price of gold is set on the London market by the five members of the London Gold Pool. It is designed to fix a price for settling contracts between members of	
RestFulServices	DOC	Try It GeoIP Service	Provide location detail for a given IP address.	
RestFulServices	DOC	Try It FedACH Service	The FedACH Services product suite provides financial institutions with efficient, low-cost batched payment services that enable an electronic exchange of debit and credit transactions through the Automated Clearing House (ACH) network	
RestFulServices	DOC	Try It Bible - King James Version	Pro	
RestFulServices	DOC	Try It Currency Converter	Pro	
RestFulServices	DOC	Try It Stock Quote	Pro	
robertlie	RPC	Try It Countries web service	Ge	
SoaMoa	RPC	Try It SoaMoa Sample Services	countryname, regionname, latitude, longitude)	PHP
			A sample artist registry webservice	JAX-RPC

Dahinter: Fast immer ein DBMS!

Beispiele: Anwendung

45

SAP Display Standard Order 117: Overview

Document flow | Status overview | Sales summary | Sold-to party | Orders

Sales area: 0102 / 10 / 20 Sales grp: Sales off.

Sold-to party: 100578 HOUGHTON CHEMICAL CORP // ALLSTON MA 02134

Ship-to party: 100578002 HOUGHTON CHEMICAL CORP / 30 AMOR AVENUE / CARLSTADT ...

Standard Order: 117 Net value: 34,463.62 USD

PO number: PO date:

Sales | Item overview | Item detail | Ordering party | Procurement | Shipping | Reason for rejection

General header data

Sales doc.type: OR Standard Order Delivery prior.: 2

Contact person: ATTN: QUALITY Cont. tel.:

Req. deliv.date: D 09/24/2002 Total weight: 45,311.375 KG

Payment terms: L030 NET 30 DAYS INVO... Incoterms: FOB Destination

Shp.Cond.: 90 CPU or Arranged Title Transfer

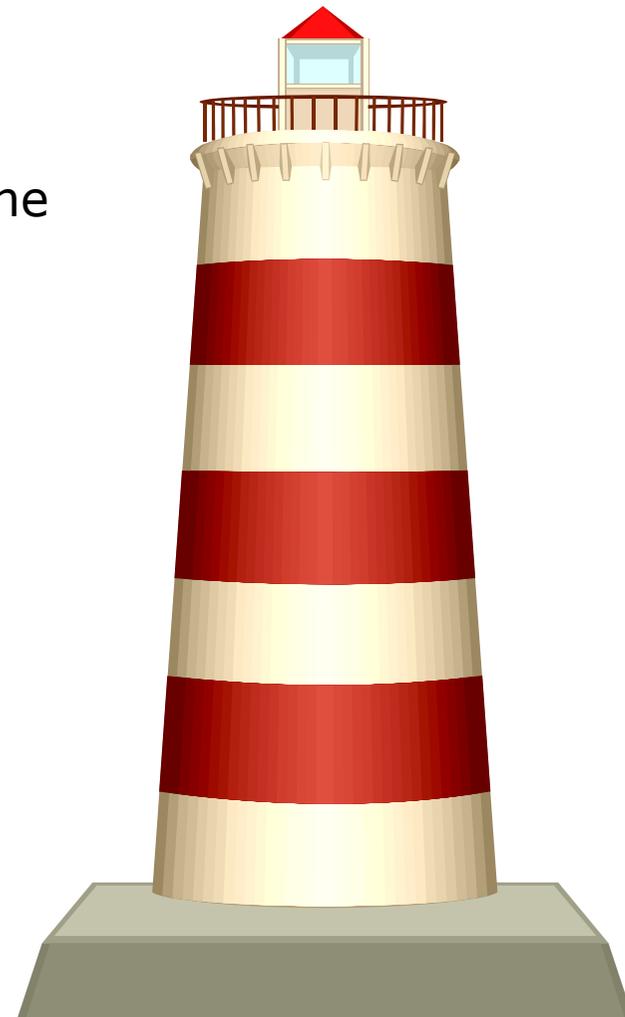
All items

Item	Material	Description	Customer material no.	Order quantity	SU	First date	S	PO details
10	add 5017236	LUENE BT BULK ST...		2,500.000	GAL	09/24/2002		otc 082
20	add 5019003	XYLENE BT BULK STD...		4,000.000	GAL	09/24/2002		otc 082
30	add 5017236	TOLUENE BT BULK ST...		45,000	LB	09/25/2002		otc 082

Dahinter: Fast immer ein DBMS!

VA03 hoeaag71 INS

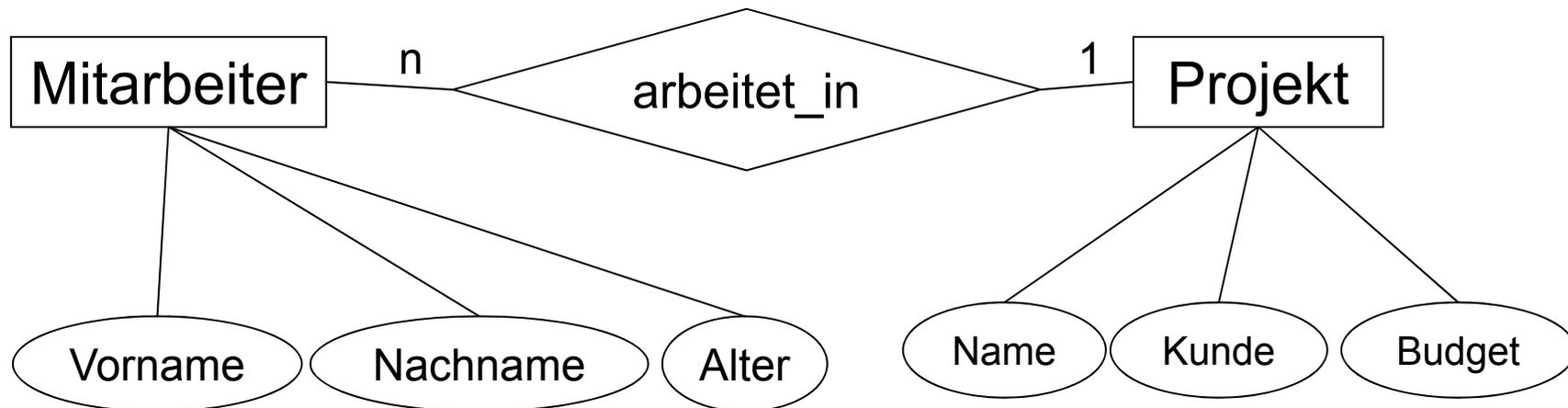
- Vorstellung der Arbeitsgruppe
- Organisatorisches
- Datenbanken und Informationssysteme
- Das Semester an einem Beispiel
- Ausblick auf das Semester



Datenbankentwurf – ER Diagramme

47

Mitarbeiter und Projekte



Nicht in dieser Vorlesung – bekannt aus Modellierung I

Relationales Modell – Relationen

48

mitarbeiter				
p_id	vorname	nachname	alter	proj_id

```
CREATE TABLE mitarbeiter(
    p_id          INTEGER,
    vorname       CHAR(25),
    nachname      CHAR(50),
    alter         INTEGER,
    proj_id       INTEGER,
    PRIMARY KEY(p_id),
    FOREIGN KEY proj_id
    REFERENCES projekte )
```

projekte			
proj_id	name	kunde	budget

Relationales Modell – Extension

49

mitarbeiter

p_id	vorname	nachname	alter	proj_id
1	Peter	Müller	32	1
2	Stefanie	Meier	34	1
5	Petra	Weger	28	2
7	Andreas	Zwickel	44	5
...	

projekte

proj_id	name	kunde	budget
1	DWH	BMW	400000
2	eCommerce	Metro	20000
5	SAP	RAG	50000
...

Integrität

50

- Datentypen
- Schlüssel
- Fremdschlüssel
- Rechte
- Transaktionen

mitarbeiter				
p_id	vorname	nachname	alter	proj_id
1	Peter	Müller	32	1
2	Stefanie	Meier	34	1
5	Petra	Weger	28	2
7	Andreas	Zwickel	44	5
...	

projekte			
proj_id	name	kunde	budget
1	DWH	BMW	400000
2	eCommerce	Metro	20000
5	SAP	RAG	50000
...

Anfragen

51

- Deklarativ:
 - Nicht „Wie erzeuge ich ein Anfrageergebnis?“
 - Sondern „Was soll im Anfrageergebnis stehen?“
- Sprachlich
 - Nachnamen aller Personen, die in kleinen Projekten arbeiten
- Relationale Algebra
 - $\Pi_{m.nachn.} \sigma_{p.budget < 40000} (\sigma_{m.proj_id = p.proj_id}(\text{mit. } x \text{ proj.}))$
- SQL
 - **SELECT** m.nachname
FROM Mitarbeiter m, Projekte p
WHERE p.budget < 40000
AND m.proj_id = p.proj_id

Anfragebearbeitung – Problem

52

Anfragen

- Anfragen sind deklarativ.
- Anfragen müssen in eine ausführbare (prozedurale) Form transformiert werden.

Ziele

- QEP – prozeduraler Query Execution Plan
- Optimierung (Effizienz)
 - Schnell
 - Wenig Ressourcenverbrauch (CPU, I/O, RAM, Bandbreite)
 - Green IT: Stromverbrauch senken.

Anfragebearbeitung

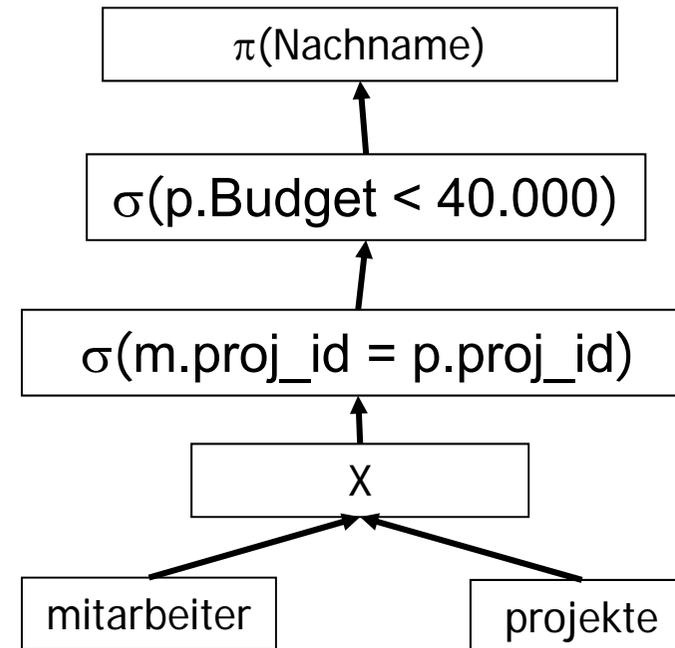
53

- Genereller Ablauf
 1. Parsen der Anfrage (Syntax)
 2. Überprüfen der Elemente (Semantik)
 3. Berechnung von möglichen Ausführungsplänen
 - ◇ Exponentiell viele
 4. Wahl des optimalen Ausführungsplans
 - ◇ Regelbasierter Optimierer
 - ◇ Kostenbasierter Optimierer
 5. Anfrageausführung

Anfragebearbeitung – Ausführung

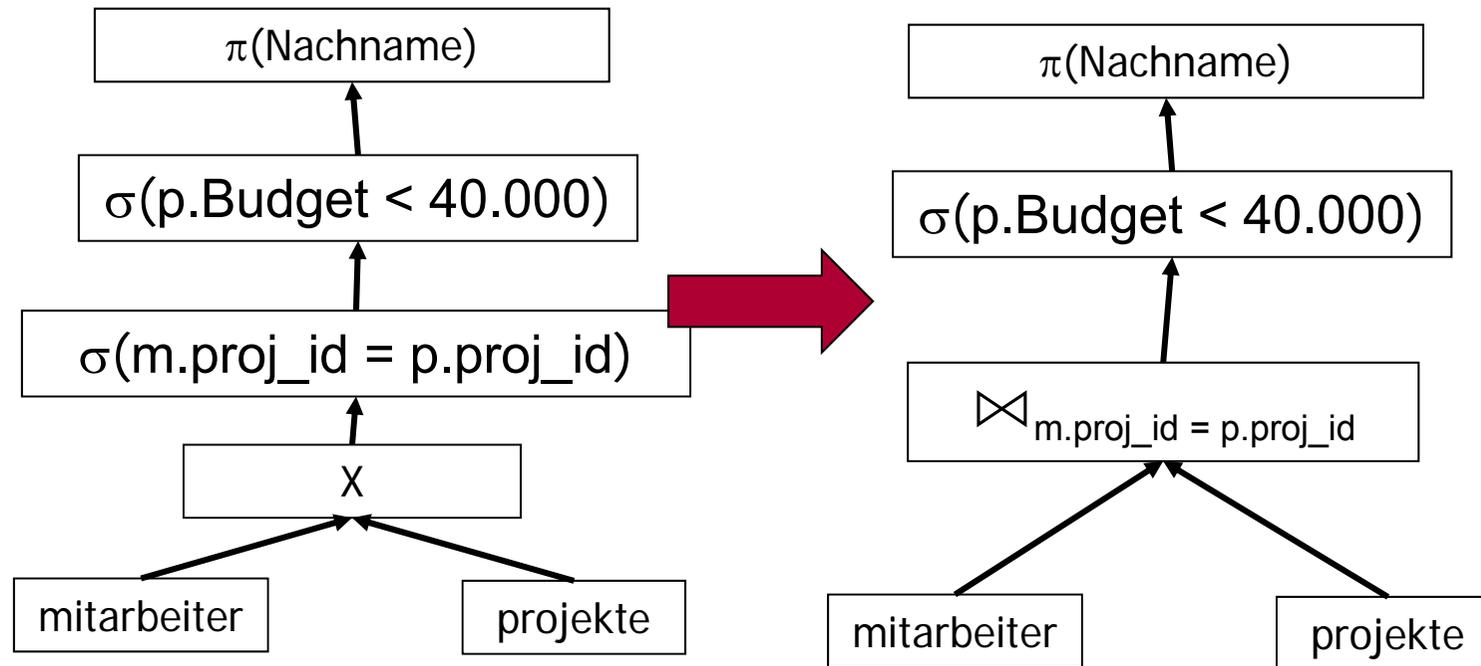
54

- Interpretation des Ausführungsplans als Baum
 1. Kartesisches Produkt
 2. Zwei Selektionen
 3. Projektion
- Sehr aufwendig!



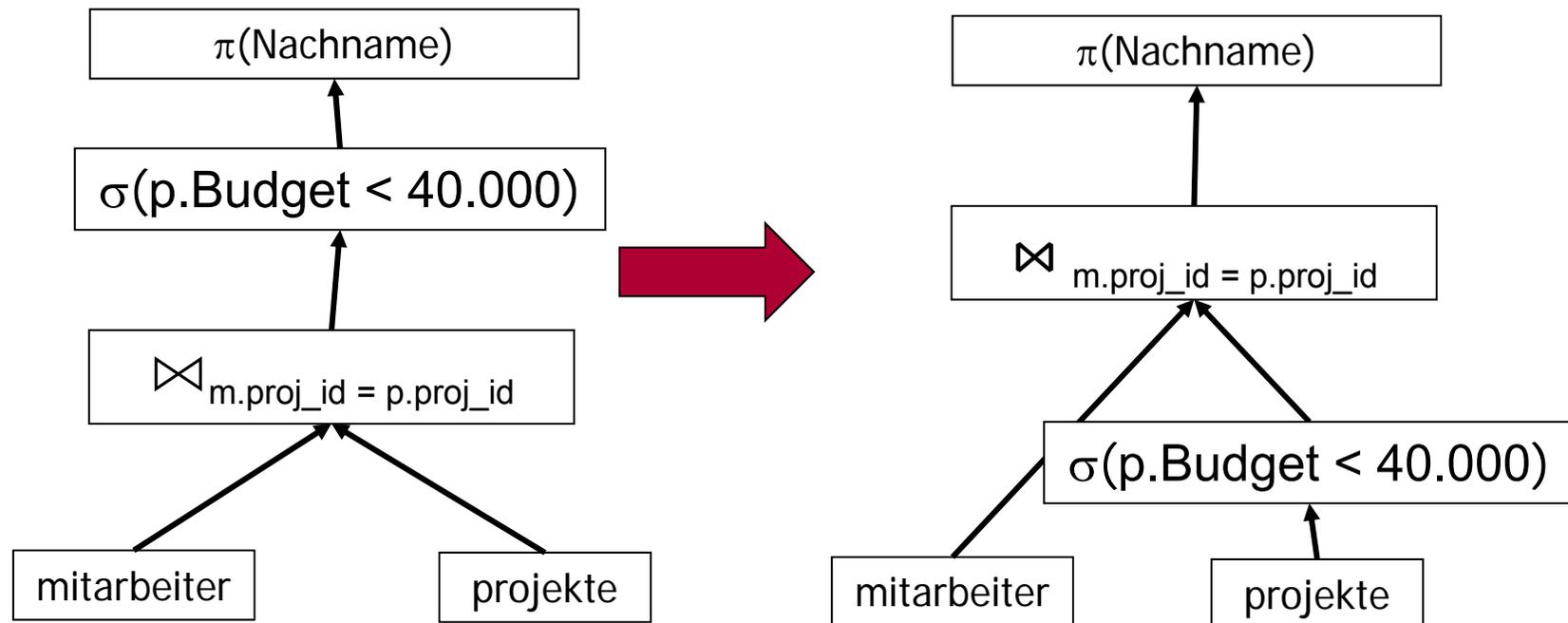
Anfragebearbeitung – Beispiel

55



Anfragebearbeitung – Beispiel

56



Sichten

57

- **CREATE VIEW** MA_MiniProjekte AS (
 SELECT m.nachname
 FROM Mitarbeiter m, Projekte p
 WHERE p.budget < 40000
 AND m.proj_id = p.proj_id)

- **SELECT * FROM** MA_MiniProjekte
- Probleme:
 - Anfrageplanung
 - Materialisierung
 - Updates durch Sichten hindurch

Transaktionsmanagement

58

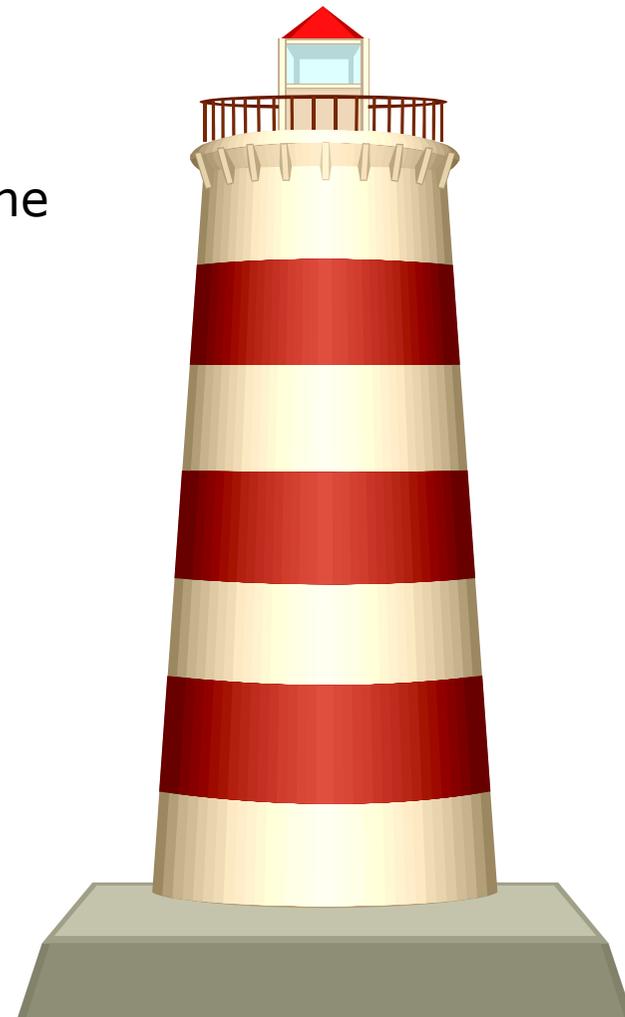
- Problem: Mehrbenutzerbetrieb
 - **INSERT INTO** mitarbeiter **VALUES** (Hans, Müller, 36, 5)
 - **DELETE FROM** projekte **WHERE** proj_id = 5

- Reihenfolge der Operationen ist nicht unabhängig
 - Serialisierbarkeit
 - Sperren

- XML und XML Schema
- Speicherung datenzentrierter XML-Dokumente
- Speicherung dokumentzentrierter XML-Dokumenten
- Speicherung von XML in bestehende Datenbank
- XPath und XQuery
- Indizierung



- Vorstellung der Arbeitsgruppe
- Organisatorisches
- Datenbanken und Informationssysteme
- Das Semester an einem Beispiel
- Ausblick auf das Semester



Kurzüberblick

61

- Einführung & Beispiel
 - Datenbanksysteme - Historie →
 - Relationaler Datenbankentwurf →
 - Relationale Algebra →
 - SQL →
 - Anfragebearbeitung und -optimierung →
 - Transaktionsmanagement →
 - XML →
 - Data Warehouses
- Übung 1: Historische DBMS
 - Übung 2: Datenbankentwurf
 - Übung 3: Relationale Algebra
 - Übung 4: SQL
 - Übung 5: JDBC
 - Übung 6: Transaktionen
 - Übung 6: XML
 - Übung: Klausurvorbereitung

Die Themen im Einzelnen

62

- Was sind Datenbanken?
 - Motivation, Historie, Datenunabhängigkeit, Einsatzgebiete
- Datenbankentwurf im ER-Modell
 - Entities, Relationships, Kardinalitäten, Diagramme
- Relationaler Datenbankentwurf
 - Relationales Modell, ER -> Relational, Normalformen, Transformationseigenschaften
- Relationale Algebra
 - Kriterien für Anfragesprachen, Operatoren, Transformationen
- SQL
 - SQL DDL, SQL DML, SELECT ... FROM ... WHERE ...

Die Themen

63

- Integrität und Trigger
 - Fremdschlüssel, weitere Bedingungen, Trigger
- Datenbankprogrammierung
 - JDBC
- Sichten und Zugriffskontrolle
 - Sichtenkonzept, Änderungen auf Sichten, Rechtevergabe
- Anfragebearbeitung und -optimierung
 - Anfragebäume, Kostenmodell, Transformationen, Dynamische Programmierung
- Transaktionsverwaltung
 - Mehrbenutzerbetrieb, Serialisierbarkeit, Sperrprotokolle, Fehlerbehandlung, Isolationsebenen in SQL
- XML
 - Datenmodell, XPath, XQuery, Speicherung, Indizierung
- Ausblicke
 - Data Warehouses, Data Mining, Informationsintegration, DBS II

- Fragen bitte jederzeit!
 - In der VL
 - Sprechstunde: Dienstags 14 - 15 Uhr
 - Email: naumann@hpi.uni-potsdam.de
 - Telefon: (0331) 5509 280

- Anregungen zur Verbesserung!

The end